第 37 卷第 1 期
录用定稿
计算机应用研究
Application Research of Computers
Vol. 37 No. 1
Accepted Paper

一种基于改进网络编码的匿名通信机制研究 *

杨　康，翟江涛[†]，戴跃伟

(江苏科技大学 电子信息学院，江苏 镇江 212003)

摘　要： 现有的基于网络编码的匿名系统，由于转发节点的不稳定性以及匿名消息的多路径转发，使得数据发送的成功率较低，提出了一种基于网络编码与数据冗余方法的新的匿名通信机制 ACSNC（anonymous communication system based on network coding）。首先采用数据冗余机制对要发送的数据进行冗余分片并沿多条路径发送；然后中间转发节点对信息进行随机编码后转发，使信息通过节点前后的统计特性发生变化；最后，目的节点根据收到的数据片与编码系数恢复匿名信息。仿真结果表明，该机制在显著提高数据发送成功率的基础上，能够有效地保障通信的匿名性与安全性。

关键词： 匿名通信；网络编码；数据冗余；多路径；网络安全

中图分类号： TN918　　doi: 10.19734/j.issn.1001-3695.2018.06.0475

Anonymous communication mechanism based on improved network coding

Yang Kang, Zhai Jiangtao[†], Dai Yuewei

(*School of Electronics & Information Jiangsu University of Science & Technology*, *Zhenjiang Jiangsu* 212003, *China*)

Abstract: In the existing anonymous system based on network coding, the success rate of data transmission is low due to the instability of forwarding nodes and the multipath forwarding of anonymous messages. Aiming at this, a new anonymous communication mechanism ACSNC based on network coding and data redundancy method is proposed. The proposed mechanism first adopts a data redundancy mechanism to redundantly fragment the data to be sent and send it along multiple paths. Then, the intermediate forwarding node randomly encodes the information and then forwards the information so that the statistical characteristics of the information are changed before and after passing through the node. Finally, the destination node recovers the anonymous information based on the received data slice and coding coefficients. Simulation results show that the mechanism can effectively guarantee the anonymity and security of communication on the basis of significantly improving the success rate of data transmission.

Key words: anonymous communication; network coding; data redundancy; multipath; network security

0　引言

无处不在的网络监控、流量审查给互联网中的隐蔽通信带来严重威胁。在敏感数据传输时，人们不仅希望数据的透明传输，还希望通信双方能够匿名。然而，网络通信本身并不提供对身份信息的保护。因此，匿名通信机制成为身份信息保护的研究热点[1]。

匿名通信是指通信过程中通过某种技术手段保护通信双方之间的关系，使得攻击者不能直接获得或间接推断出双方之间的通信关系或任何通信一方的身份信息[2]。Mix-Net 技术起源于 1981 年由 David Chaum[3]提出，自此之后的三十多年来，众多学者坚持不懈的讨论和发展匿名通信技术，经典的研究成果不断涌现。国外著名的匿名研究机构 Free Haven[4]收录的匿名通信领域的文章逐年递增，其中有很多值得分析和借鉴，如 Anonymizer[5]、洋葱路由[6]、Tor[7]、Crowds[8]等。然而传统的匿名通信系统一般都是基于可信的密钥基础设施，需要通过第三方参与密钥的协商，这就会导致整个通信系统的安全性很大的依赖于密钥的安全性。此外，传统的匿名系统也很难适应无密钥信任中心的网络环境。

因此，对于无密钥中心的匿名通信研究具有重要的意义。

收稿日期： 2018-06-19；**修回日期：** 2018-07-24　　**基金项目：** 国家自然科学基金资助项目（61702235，61472188，61602247，U1636117）；江苏省自然科学基金资助项目（BK20150472，BK20160840）；国家科技支撑计划资助项目（2014BAH41B01）；中央高校基本科研业务费专项资金资助项目（30920140121006，30915012208）

作者简介： 杨康（1993-），男，江苏泗阳人，硕士研究生，主要研究方向为多媒体与信息安全、；翟江涛+（1983-），男（通信作者），河南三门峡人，副教授，博士，主要研究方向为多媒体与信息安全（jiangtaozhai@gmail.com）；戴跃伟（1962-），男，江苏镇江人，教授，博导，主要研究方向为多媒体与信息安全、系统工程理论及应用、复杂系统管理控制.

文献[9]通过源信息分割策略，实现无密钥中心下的匿名路径构建。分析表明该文所述的匿名系统在路径构建阶段显著的提高了系统的抗攻击性，但是该系统在信息传输阶段没有使用编码技术来对匿名信息进行编码处理。文献[10]在此基础上，进一步提出了通过信息分割以及网络编码来传输匿名通信信息。但是文献中转发节点对于匿名消息的编码转发增加了系统的延时性，同时匿名消息的多路径发送降低了匿名消息发送的成功率。针对上述问题，本文提出的匿名通信机制结合了目前网络编码的优势，使系统不需要依赖密钥机制，减少系统部署的代价。同时，引入的数据冗余机制，提高了匿名信息发送的成功率，提升了系统的效率与稳定性。

1　架构与方法

1.1　本文架构

网络编码[11,12]是一种融合了路由和编码的信息交换技术，它的核心思想是在网络中的各个节点上对各条信道上收到的信息进行线性或者非线性的处理，然后转发给下游节点，中间节点扮演着编码器或信号处理器的角色。Ahlswede 等人以蝴蝶网络[13]的研究为例指出，通过网络编码可以达到多播路由传输的最大流界，提高了信息的传输效率。网络编码技术可以提高现有的网络吞吐量，同时还能改善网络的可靠性和防范攻击的能力。在匿名通信中，通过中间节点对传输的信息进行编码处理，使得数据流在流入和流出节点的形式与内容发生了变化。使得攻击者无法进行信息流的追踪。但是现有的基于网络编码在多播的场景下缺乏很难保证系统数据发送的成功率，针对此，本文通过引入数据冗余机制改进了基于网络编码的匿名通行机制。冗余分片的机制如图 2 所示。读取文件进行分片，得到 N 个分片，然后对每一行的分片进行异或运算，结果保存到辅助分片中，对每一列也进行异或运算，结果也保存到辅助分片中，结果得到 $2N+1$ 个分片，随机删除 x 个分片 $(1<x<N+1)$,即可能删除原有分片，也可能删除辅助分片，然后利用剩下的分片即能恢复原来的数据[14]。

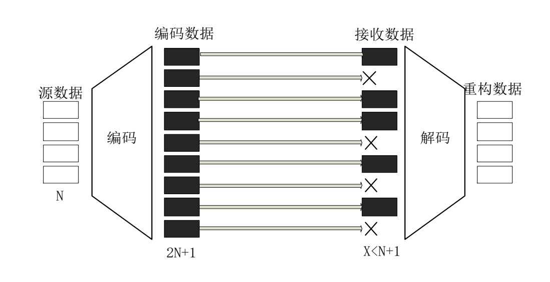


图 1　数据冗余分片机制的过程

Fig.1　Process of data redundancy fragmentation mechanism

1.2　设计思想

首先，匿名消息的源发送节点通过与目录服务器的交互，采取可信计算来选取可信任的节点构建匿名路径的转发网络[15,16]。确保每个转发节点仅知道其前趋节点与后继节点。在建立匿名路径后，源发送节点通过数据冗余机制，将消息分片并通过多个路径发送。转发节点在接收到匿名数据后，通过随机编码系数对数据包进行编码转发，使数据包在流入和流出节点时，统计特性发生变化。同时，将随机编码系数和编码后的数据一起继续转发给下一个节点直到目的节点。最终，接收者收到所有的消息分片以及随机编码系数，通过相应的解码规则还原匿名消息。系统的架构如图 2 所示。

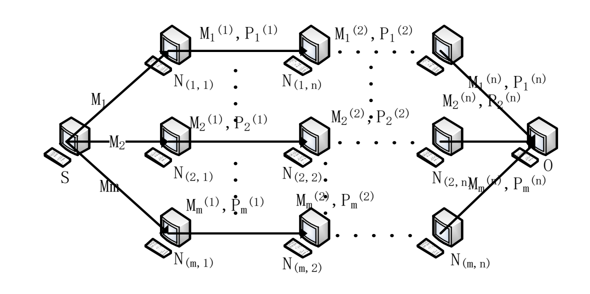


图 2　匿名编码通信网络

Fig.2　Anonymous coded communication network

1.3　改进方法

本节将详细叙述 ACSNC 机制的通信过程，表 1 列出了系统中常用符号的定义。

表 1　系统中符号定义

Table 1　Symbol definition in the system

符号	含义
S	源发送节点
O	目的节点
M_m	分片消息
$N_{(m,n)}$	转发节点
$P_m^{(n)}$	编码系数

1）匿名转发网络建立

匿名网络中转发节点的路由构建主要借鉴文献[9]中基于网络编码的匿名建路机制。但文献[9]中转发节点对编码系数的多路径转发，增加了系统的复杂性和时延性。本文改进了转发节点的转发机制，将转发系统通过当前路径转发，减小了时延，同时提升了系统的容错率。最终，实现每个转发节点仅能够知晓其下一个节点的地址。

如图 1 所示，源发送节点为 S，目的节点为 O，转发节点为 $N_{(m,n)}$。如果匿名系统需要构建一条 $S-N_1-N_2-N_3-O$ 的链路。则源节点 S 向 N_1 发送建路信息时，通过与目录服务器的交互中随机选取除了建路节点 $N_1-N_2-N_3$ 节点外的 $m*n$ 个节点构建编码转发节点。之后，源节点将建路信息编码转发到目的节点 N_1。其他节点的路由信息构建重复上述步骤即可。最终，在整个编码转发网络中每个转发节点将获取其后继节点的 ip。

2）匿名消息转发阶段

a)源发送节点首先对匿名数据进行冗余处理，冗余编码的模型如图 3 所示。先将匿名数据 M 拆分成 $k(k<m)$ 个数据包。然后，通过异或算法，生成 1 或多个冗余数据包。接着，将切

分的 $m_1...m_k$ 个数据和生成的冗余数据 R，通过 m 条不同的链路传输。

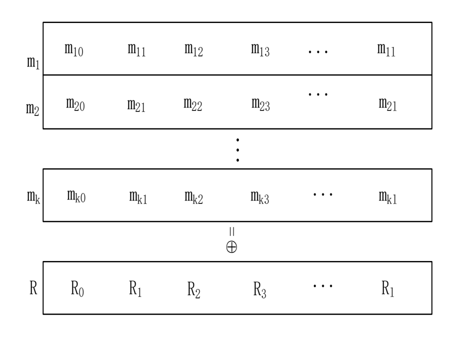


图 3　冗余编码模型

Fig.3　Redundant coding model

b)匿名链路中的转发节点 $N_{(m,n)}$ 在收到前一个节点发送来的匿名数据后，选择随机编码系数 $P_m^{(n)}$ 对接受到的匿名数据 $M_m^{(n)}$ 进行编码处理，之后将编码后的数据 $M_m^{(n+1)}$ 以及随机编码系数 $P_m^{(n)}$ 一起发送给下一个节点 $N_{(m,n+1)}$。

c)目的节点接受到所有或者部分分片数据 M 以及随机编码系数 P。先通过获取的所有编码系数，来解码各个消息分片。最后，通过冗余算法来解码出原匿名数据消息。

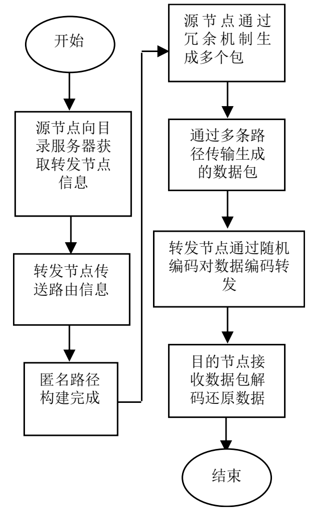


图 4　匿名通信流程

Fig.4　Anonymous communication process

2　实验与分析

2.1　匿名性分析

1）身份匿名

系统中的每个节点仅知道其前驱节点与后继节点，即使一个节点是恶意节点，通过其观察其数据包的 *ip* 也不能够获取到系统的源发送节点与目的节点的真实地址。此外，通过中间节点的编码转发，使得本系统的链路具有不可追踪性，进一步的保障了系统中源节点与目的节点的真实地址。

2）通信匿名

源节点通过冗余机制，对通信消息进行分片后通过不同路径传输，攻击者很难控制所有路径上的节点，来还原出通信信息。同时，每个转发节点对通信消息进行编码转发，进一步阻止流量分析的攻击，有效的保障了系统通信的匿名性。

3）匿名度仿真

定义匿名度 D 为分片消息在传输时，被攻击者合谋解码的概率。当匿名度 D 的值越小时，系统的匿名度越强[17]。

在编码转发网络中，攻击者想要获取通信消息，必须控制 m 条链路中每条链路上的一个节点并者这 m 个节点进行数据共享，才能解码还原出源通信消息。

攻击者通过合谋攻击获得通信消息的概率为 $\frac{n^m}{C_{mn}^m}$，即匿名度 $D=\frac{n^m}{C_{mn}^m}$。其中 m，n 分别表示匿名转发网络中的链路数与单条链路的节点数。如图 5，6 所示，ACSNC 的匿名度随着 n 与 m 的增大而减小，但受传输的链路数 m 的影响较大。从图 6 可以看出，当 $m\geq6$，$n\geq3$ 时，系统的匿名度 D 的值趋近于 0，此时，接近于绝对匿名，同时表明再通过增加链路数与转发的节点数来提升系统的匿名性效果不明显。信息的泄露主要由于攻击者对于转发节点的控制，匿名消息被分成的分片数越多，沿着更多的路径传输，攻击者就难以控制每条转发路径上的一个节点，系统的匿名性就会更强。同时，增加单条路径上的转发节点数，使得匿名路径的长度变长，进一步增强了系统的匿名性。

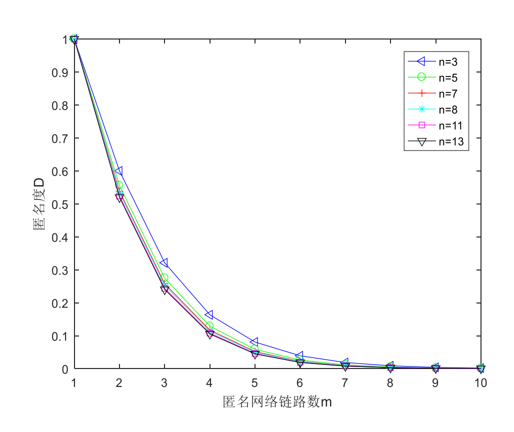


图 5　n 对匿名度的影响

Fig.5　The impact of n on anonymity

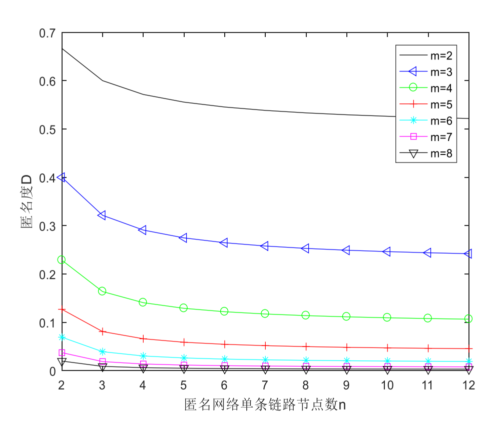


图 6　m 对匿名度的影响

Fig.6　The impact of m on anonymity

2.2 性能分析

在源节点发送信息时，ACSNC 通过数据冗余机制，对信息进行冗余分片处理。通过冗余分片的机制，可以提高系统数据的发送成功率。通过对原始数据增加一个或多个冗余包，使得在 m 条链路上丢失一个或多个数据包后，接收端仍能还原原数据。假设每条链路上的传输成功率相同为 η^n，则未加冗余包时，m 条链路发送成功的概率如公式(1)，增加单一的冗余包后的成功发送概率公式如(2)。生成一个冗余包的仿真结果如下图(7)。同时，ACSNC 引入的冗余机制也支持生成多包冗余，其仿真结果如图 8 所示。

$$P_1 = \eta^n \tag{1}$$

$$P_2 = \eta^{n+1} + C_{n+1}^{n}(1-\eta)\eta^n \tag{2}$$

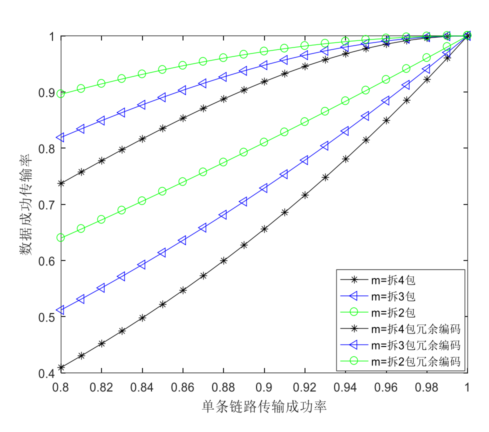


图 7 生成单个冗余包的成功传输率

Fig.7 Successful transfer rate for generating a single redundant package

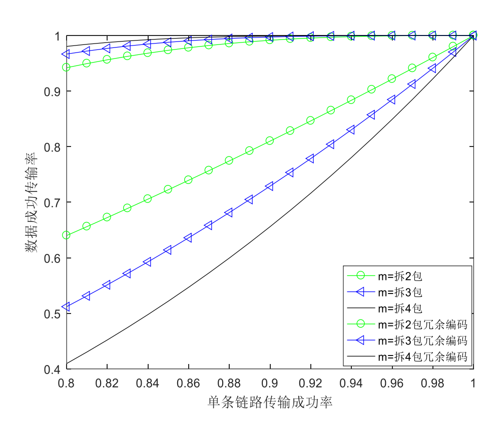


图 8 生成多个冗余包的成功传输率

Fig.8 Successful transfer rate for generating multiple redundant packets

由图 7、8 可以看出，通过冗余机制来发送数据包对于数据发送的成功率提高效果显著。图 7 中通过生成单个冗余包，可以看出系统的数据发送成功率提升了 20%。图 8 通过生成多个冗余包其提升效果显著，提升了 30 以上的发送成功率。因此，通过在匿名通信中引入冗余分片机制，极大的提高了系统的稳定性和容错率。

2.3 安全性分析

1）合谋攻击

源节点首先通过冗余机制，将消息分成多片通过多条路径传输。同时，系统中的每个转发节点，仅仅知晓其前驱接节点与后继节点的地址。攻击者想要通过合谋攻击来探知整个系统的匿名通信关系很困难。同时通过 2.1 的匿名分析可以得到，只要控制好系统中的链路数以及转发节点数，攻击者攻击成功的概率基本可以忽略。

2）流量分析攻击

ACSNC 的转发网络中，每个中间转发节点都对数据进行编码转发，使得传输的信息在流入与流出节点时，数据的形式发生了变化。使得攻击者想要通过流量分析来追踪数据包传送的路径，从而探知到源节点与目的节点的地址变的相当困难。

3 结束语

本文基于网络编码思想，通过引入数据冗余机制，提出了改进网络编码匿名通信机制 ACSNC。ACSNC 承继了现有的基于网络编码的匿名系统优势，同时通过引入的数据冗余机制，极大的提高了匿名数据发送的成功率。通过理论分析和仿真结果表明，ACSNC 具有较好的稳定性、匿名性和安全性。本文是在假设各个转发节点的转发概率相同的情况下进行的研究，未来将考虑更复杂的网络环境下系统的效率性。

参考文献：

[1] Haraty R A, Assi M, Rahal I. A systematic review of anonymous communication systems [C]// Proc of International Conference on Enterprise Information Systems. [S. l.] : SciTePress, 2017: 211-220.

[2] Zhu Tingwei, Feng Dan, Wang Fang, *et al.* Efficient anonymous communication in SDN-Based data center networks [J]. IEEE//ACM Trans on Networking, 2017, PP (99): 1-14.

[3] Chaum D. Untraceble electronic mail, return addresses, and digital pseudonyms [J]. Communications of the ACM, 1981, 24 (2): 84-90.

[4] Fanti G, Viswanath P. Algorithmic advances in anonymous communication over network [C]// Proc ofConference on Information Science and Systems. Piscataway, NJ: IEEE Press 2016: 12-21.

[5] Justin B. The anonymizer: protecting user privacy on the Web [J]. Computer-Mediated Communication Magazine, 1997, 4 (9): 34-38.

[6] Reed M G, Syverson P F, Goldschlag D M. Anonymous connections and onion routing [C]// Selected Areas in Communications, Piscataway, NJ: IEEE Press, 1998, 16 (4): 482-494.

[7] Dingledine R, Mathewson N, Syverson P. Tor: the second-generation onion router [J]. Journal of the Franklin Institute, 2004, 239 (2): 135-139.

[8] Reiter M K, Rubin A D. Anonymous Web transcations with crowds [J]. Communications of the ACM, 1999, 42 (2): 32-48.

[9] 段桂华，王伟平，王建新，等. 一种基于多路径网络编码的匿名通信机制 [J]. 软件学报, 2010, 21 (9): 2338-2351. (Duan Guihua, Wang Weiping, Wang Jianxin, Yang Luming. Anonymous communication mechanism with Multi-Paths network coding [J]. Journal of Software, 2010, 21 (9): 2338-2351.)

[10] 周彦伟，杨波，吴振强，等. 基于网络编码的匿名通信模型 [J]. 中国



科学：信息科学, 2014, 44 (12): 1560-1579. (Zhou Yanwei, Yang Bo, Wu Zhenqiang, *et al.* Anonymous communication model based on network coding [J]. Chinese Science: Information Science, 2014, 44 (12): 1560-1579.)

[11] Chen Fei, Xiang Tao, Yang Yuanyuan, *et al.* Securecloud storage meets with secure network coding [J]. IEEE Trans on Computers, 2016, 65 (6): 1936-1948.

[12] Yu Mingchao, Sadeghi P, Aboutorab N. Performance characterization and transmission schemes for instantly decodable network coding in wireless broadcast [J]. Eurasip Journal on Advances in Signal Processing, 2015, 2015 (1): 1987-1996.

[13] Ahlswede R, Cai N, Yeung R W, *et al.* Network information flow [J]. IEEE Trans on Information Theory, 2002, 46 (4): 1204-1216.

[14] 王少辉, 蒋季宏, 肖甫. 基于重路由匿名通信系统的设计 [J]. 计算机科学, 2016, 43 (10): 154-159. (Wang Shaohui, Jiang Jihong, Xiao Wei. New design of rerouting-based anonymous communication system [J]. Computer Science, 2016, 43 (10): 154-159.)

[15] 谭庆丰, 时金桥, 方滨兴, 等. 匿名通信系统不可观测性度量方 [J]. 计算机研究与发展, 2015, 52 (10): 2373-2381. (Tan Qingfeng, Shi Jinqiao, Fang Binxing, *et al.* Towards measuring unobservability in anonymous communication systems [J]. Computer Research and Development, 2015, 52 (10): 2373-2381.)

[16] 周彦伟, 杨波, 张文政. 新型的多路径匿名通信系统 [J]. 电子学报, 2017, 45 (5): 1234-1239. (Zhou Yanwei, Yang Bo, Zhang Wenzheng. A new anonymous communication system with multi-path [J]. Chinese Journal of Electronics, 2017, 45 (5): 1234-1239.)

[17] Bagai R, Malik N, Jadliwala M. Measuring anonymity of pseudonymized data after probabilistic background attacks [J]. IEEE Trans on Information Forensics & Security, 2017, PP (99): 1.